ERCAP Request #83553 for FY2010

1. Principal Investigator

Name	Phone	Email
Bethel, Edward (Wes)	510-486-7353	ewbethel@lbl.gov

2. Other Authorized Preparers

3. Senior Investigators	(Utah) Chris Johnson, Chuck Hansen, Valerio Pascucci, Claudio Silva, Allen Sanderson (LLNL) Eric Brugger, Brad Whitlock (ORNL) Sean Ahern, Jeremy Meredith, Dave Pugmire, George Oustrochov (UCD) Ken Joy, Christoph Garth, Eduard Deines (LBNL) Hank Childs, Mark Howison, Prabhat, Gunther Weber
4. Project Title	SciDAC2 Visualization and Analytics Center for Enabling Technologies
5. Project Name	VACET
6. Project Class	SciDAC
7. Sponsoring Site	USA: Lawrence Berkeley National Laboratory
8. Science Category	Computer Sciences
9. DOE Office and Program	ASCR - Computer Sciences

10. Is this project funded by the DOE Office of Science?

X
Yes, this project has direct grant support from the DOE Office of Science

Who is your DOE Program Manager?

Osni Marques and Dan Hitchcock

DOE Office of Science Grant Number(s)

1. SciDAC2 Visualization and Analytics Center for Enabling Technologies (VACET). Total award amount is $2.2M/yr for five years.
2. High Performance Visualization, ASCR/MICS Base

Program. Award is $600K yr through FY11.

According to LBL Finance people, there are no "Grant Numbers" associated with either of these awards but both are "on the books."

11.1 Does this project make use of national security information?

	Yes	X	No

11.2 Does this project collect and store Social Security Numbers, Personally Identifiable Health Information (names or other identifiers matched with health data), Driver's License Numbers, or Financial Account Numbers?

	Yes	X	No

11.3 If this project is given a NERSC award, I agree to monitor the usage associated with it to ensure that, to the best of my ability to determine, usage is for the project described here.

X	Yes		No

11.4 For continuing projects: I have audited the MPP and/or HPSS usage associated with this project, and to the best of my ability to determine, all usage was for the project specified.

X	Yes		No

12. Computational Resources Requested

Center	Resource	Alloc Type	Repo	Hours Used 2009	Hours Requested 2010
NERSC	**MPP (MPP Charging)**	DOE Production	m636	449,890	900,000

13. Mass Storage Resources Requested

Center	Resource	Alloc Type	Repo	SRUs Used 2009	SRUs Requested 2010
NERSC	**HPSS Charging**	DOE Production	m636	33,020	250,000

14. Justification for Resources Requested

Our request represents estimates needed for four different research thrusts, three of which will conduct scalability studies on franklin and will involve many 32k-way parallel runs, as well as runs at lesser concurrency.

Specifically:
1. Performance and scalability of Visualization algorithms: multiple 32K-way parallel runs, 300K hours.
2. VisIt hero runs and scalability studies: multiple 32K-way parallel runs, 300K hours.
3. Parallel Integral Curves (streamlines and stream surfaces): multiple 32K-way parallel runs, 200K hours.
4. Topological analysis: 10K hours (but a substantial

NGF requirement)

15.1 PDSF Repos

15.2 How many total CPU hours do you need for the coming year (please describe your codes on the codes tab)?

Hours

15.3 Describe any special characteristics of your workflow (e.g. non-conforming)

15.4 How much experiment specific storage do you need?

Terabytes

15.5 PDSF uses a buy-in model. How does your project plan to fund your PDSF usage (compute and storage)?

16. Project Description

16.1 Project Overview: provide a non-technical brief overview. What will this project accomplish? What is the significance of this work?

Launched in 2006 as one of nine centers under the Department of Energy's Scientific Discovery through Advanced Computing (SciDAC-2), VACET focuses on leveraging scientific visualization and analytics software technology as an enabling technology for increasing scientific productivity and insight. Advances in computational technology have resulted in an information big bang, which in turn has created a significant data understanding challenge. This challenge is widely acknowledged to be one of the primary bottlenecks in contemporary science. The vision for our Center is to respond directly to that challenge by adapting, extending, creating when necessary and deploying visualization and data understanding technologies for our science stakeholders at DOE's open computing facilities (NERSC/LBNL and NCCS/ORNL). Using an organizational model as a Visualization and Analytics Center for Enabling Technologies (VACET), we are well positioned to be responsive to the needs of a diverse set of scientific stakeholders, including other SciDAC projects, in a coordinated fashion using a range of visualization, mathematics, statistics, computer and computational science and data management technologies.

More specifically, VACET will provide visualization and analytics software infrastructure to support the challenging data understanding needs of SciDAC Science Applications. This infrastructure will be deployed at DOE's open computing facilities, both NERSC/LBNL and NCCS/ORNL.

16.2 Project Description: longer project descriptions may be attached as a PDF file

16.3 Performance and Scaling Information (optional, but recommended for requests of 500,000 or more hours)

Provide parallel scaling and performance information for the code(s) in your repository as a PDF file. If the repository has several codes, please choose codes that you wish to hightlight. Indicate the platform used to obtain the results, the approximate date(s) of the measurements, and whether the numbers given represent strong or weak scaling. Plotes of tables are preferable to a single raw performance number in flops. Also

please provide a primary contact email for each table/plot or set of results, since these results may be included in future NERSC publications and reports.

CGA-VisItHeroRuns-Fall2009-submission-skinny.pdf

16.4 URL for a relevant web page.

http://www.vacet.org/

17.1 Accomplishments Overview:
Summarize in non-technical terms the accomplishments made by this project in the previous year. Explain why your results are important and how the use of NERSC resources allowed you to meet your research objectives.

1. Performance and Scalability Characterization of Visualization Algorithms
Last year accomplishments.
* Implement and benchmark a stencil-based algorithm having a uniform memory access pattern: MPI, MPI+pthreads, MPI+OpenMP. Run at high levels of concurrency, up to 32K-way parallel on franklin.
* Preliminary implementation of a non-uniform memory access algorithm (ray-casting volume rendering), initial runs to 32K-way parallel on franklin.

2. VisIt Hero Runs
Ran studies on Franklin at 16K and 32K cores looking at isocontouring and volume rendering. Also looked at collective and non-collective I/O.

3. Parallel Integral Curves
Developed parallel streamline algorithm that is now deployed in VisIt and published in an SC09 paper.

4. Topological Analysis
* Moved from 2D-based analysis of flame surfaces to fully 3D based analysis of fuel consumption
* Extended analysis code to use the native AMR data hierarchy
* Implemented linked view space-time interface to interactively explore the evolution of burning cells

17.2 Accomplishments: longer accomplishment descriptions may be attached as a PDF file.

18. Relevant Publications: Enter in the order authors, title, journal

18.1 Refereed Publications: List all refereed publications in the last 12 months based on research using NERSC resour
may include publications submitted to journals but not publications in preparation. Please indicate if your article was highlighted on a journal cover.

1. Performance and Scalability Characterization of Visualization Algorithms
* Shoaib Kamil, Cy Chan, Sam Williams, Leonid Oliker, John Shalf, Mark Howison, E. Wes Bethel, Prabhat. "A Gen
Framework for Auto-tuning Stencil Computations." In CUG 2009, May 4-7, 2009, Atlanta GA, USA. Best Paper Awa
winner. (LBNL-2078E).
* E. Wes Bethel. "High Performance, Three-dimensional Bilateral Filtering." Lawrence Berkeley National Laboratory
Report LBNL-1601E, 2009.

2. VisIt Hero Runs
* "Extreme Scaling of Production Visualization Software on Diverse Architectures", Childs, Pugmire, Ahern, Whitloc
Howison, Prabhat, Weber, and Bethel, submitted to a special issue of Computer Graphics and Applications on large da
visualization (attached to this ERCAP application)

* Web writeups
o Scientific Computing, http://www.scientificcomputing.com/article-hpc-Smashing-the-Trillion-Zone-Barrier-082709
o HPC Wire,
http://www.hpcwire.com/topic/visualization/DOE-Researchers-Test-Limits-of-Visualization-Tool-47533672.html?vie

3. Parallel Integral Curves
Dave Pugmire, Hank Childs, Christoph Garth, Sean Ahern, Gunther Weber. "Scalable Computation of Streamlines on Large Datasets" In Proceedings of Supercomputing 2009

4. Topological Analysis
* G. Weber and P.-T. Bremer and J. Bell and M. Day and V. Pascucci, Feature Tracking Using {R}eeb graphs, Mathe and Visualization, Springer, to appear 2010.
* M. Day and J. Bell and P.-T. Bremer and V. Pascucci and V. Beckner and M. Lijewski, Turbulence effects on cellul burning structures in lean premixed hydrogen flames, Combustion and Flame, 156,pp. 1035-1045.
* P.-T. Bremer and G. Weber and V. Pascucci and M. Day and J. Bell, Analyzing and Tracking Burning Structures in Premixed Hydrogen Flames, IEEE Transactions on Visualization and Computer Graphics, 2009, to appear.
* P.-T. Bremer and G. H. Weber and J. Tierny and V. Pascucci and M. Day and J. B. Bell, A Topological Framework Interactive Exploration of Large Scale Turbulent Combustion, Proc. {IEEE} International Conference on e-Science, 2 appear.

18.2 Other Publications: List up to 5 other relevant publications in the last 12 months, also based on research using NE resources.

19.1 Code and Application Descriptions

Code Name	VisIt
Description	Our project will run multiple codes. VisIt is a production-quality, parallel-capable visualization application the uses a client-server architecture for displaying results of visualization to a remote client. It has been used on the BG/L system at LLNL to perform visual data analysis of a grid containing 27 billion zones -- it has proven to be sufficiently scalable on all modern platforms. In addition, it has been ported to run on Jaguar, so we reasonably expect it will run on all platforms at NERSC. Custom topological analysis codes. This family of codes, which are experimental, are run using large, time-varying input. Each timestep of data is processed in serial; parallelization occurs by running many such single-timestep analyses in parallel. Limiting factors for these codes is single-node memory size and I/O bandwidth. For that reason, we expect most of this work to be conducted on davinci. Custom implementations of visualization algorithms (stencil-based, uniform memory access; irregular memory access) implemented in both MPI-only as well as MPI-hybrid (with pthreads, OpenMP) for conducting scalability studies to better understand performance characteristics at scale on distributed memory multicore platforms.

Mathematics	For the topological analysis work, please refer to the following Siggraph 2007 and IEEE Visualization 2007 papers authored by a VACET team member: V. Pascucci, G. Scorzelli, P.-T. Bremer, and A. Mascarenhas. "Robust On-line Computation of Reeb Graphs: Simplicity and Speed" In Proceeding of SIGGRAPH 2007 (ACM Transation on graphics), pp. 58-1 to 58-9, 2007. Attila Gyulassy, Vijay Natarajan, Bernd Hamann, Valerio Pascucci. "Efficient Computation of Morse-Smale Complexes for Three-dimensional Scalar Functions" In IEEE Transactions on Visualization and Computer Graphics, Proceedings of IEEE Visualization 2007, October 2007.
Numerical Techniques	
Machines	Franklin, davinci, tesla, and when available, NCSC, N6
Percent of Allocation	100
Planned Processors	Everywhere from 1K to 32K.
Number of Processors Description	maximum possible
Languages	c++,MPI,MPI/IO,Python,UPC,c
Libraries	hdf4,hdf5,netcdf,papi,OpenGL, Xlib, Qt, Posix communication & I/O, Boxlib, Chombo, OpenRM
Performance Limits	disk space,MPI I/O,Posix I/O,Single Node Memory Performance,Need large memory footprint for those tasks that are serial
Performance Comments	We are already using PAPI for instrumenting some code.
Checkpoint?	N
Code Plans	Continued evolution of multicore strategies. Would like to expand into GPU-based multicore studies, particularly distributed memory, multiple-GPU configurations. Presently, NERSC has no such hardware like this.
Eigensolvers	fixed_structured,fixed_unstructured,adaptive

Other Algorithms

20. How much storage space do you need?

20.1 How much space do you need?

Home (permanent) space (Gigabytes):	1,024
Scratch space (Terabytes; mpp users only):	100
NERSC Global Filesystem (Terabytes):	50
HPSS archival storage (Terabytes):	50
Space needed for checkpoints (Terabytes; mpp users only):	0
Space needed for simulation output (Terabytes):	10
How much simulation data does the program typically read (Terabytes)?	10
How much data do you expect your project to move to NERSC from external sources (Terabytes)?	10

21. Data and I/O

21.1 Describe data sharing needs within NERSC (either between machines or between users on one machine). How do you share, why, how much data?

Typical use model: we want fastest possible I/O path for our data intensive codes. We will stage data as needed depending upon which machine we're using for a given task on the appropriate resource.

21.2 Describe data sharing needs between NERSC and other sites. How, which sites, how much data, why?

Typical use model: transfer stakeholder data to NERSC (if it is not already at NERSC, most of our stakeholder data is already at NERSC), where we then use it as input for our data intensive codes.

We often "share" data with science stakeholders by creating directories on NGF with appropriate permissions.

21.3 How do you perform I/O? Do you write concurrently from all processors to shared files (parallel I/O) or one file per processor or node? Have you experienced problems or constraints due to I/O?

Combination of parallel collective and non-collective I/O: it depends upon the code we're running and problem we're studying.

Note this question is stated incorrectly: when all processors write to a shared file, that is known as "collective I/O". One file per processor is still parallel I/O.

I/O times tend to dominate total wallclock execution time for large data sets; having access to the maximum amount of I/O bandwidth possible is a priority for us.

22. Analytics Applications

22.1 Which applications/libraries do you use or intend to use for data analysis and visualization at NERSC?

Code Analytic Apps (Q 22)

Python,
R,
VisIt,
several custom codes

22.2 Could your project benefit from NERSC visualization/data analysis assistance? Please explain.

We are already working closely with the NERSC Visualization team.

23. Networking

24. Other HPC Support

A modest allocation from the Director's Discretionary pool at NCCS/ORNL on both Jaguar and Lens. We also are able to get time on Ranger/TACC, and several LLNL machines (Dawn/BGP, Juno/CHAOS, Purple/AIX).

25. Additional Information

Our "VisIt hero runs" this past year received wonderful support from NERSC user services. Our understanding is that these were the first-ever vis runs of this size on franklin, and that the "magic MPI environment variable settings" needed to make this work helped blaze the trail for others at NERSC to also run at such high levels of concurrency.

Our resource request reflects that we are doing lots of scalability studies.

26. Feedback and Project Requirements

27. Current Request Status

Not Finished: This request is not yet ready for review.

X
Finished: This request is finished and ready for review.